```
    1   ATAATGGGAA  AGTAC|ATG|G  TTATTAAACC  CGTAACAACT  CCGAGTGTAA

   51   TACAATTAAC  GCCTGATGAT  AGAGTAACGC  CTGATGATAA  AGGTGAATAT

  101   CAACCCGTTG  AAAAGCAAAT  AGCGGGAGAT  ATAATACGTG  TACTAGAATT

  151   CAAGCAAACA  AATGAAAGTC  ATACAGGATT  GTATGGAATT  GCATATCGAG
                                                    TOX F2
  201   CTAAGAAAGT  AATAATAGCA  TATGCTTTAG  CGGTAAGTGG  TATTCATAAT

  251   GTCTCTCAAC  TTCCAGAAGA  CTATTATAAA  AATAAGGATA  ACACAGGTAG

  301   AATTTATCAA  GAATACATGT  CTAATCTTTT  ATCTGCACTA  TTGGGTGAGA

  351   ATGGTGATCA  AATTTCTAAA  GATATGGCAA  ATGATTTTAC  CCAGAACGAA

  401   CTGGAGTTTG  GAGGTCAACG  TCTTAAAAAT  ACCTGGGATA  TTCCTGATCT

  451   TGAGAATAAA  CTATTGGAAG  ATTATTCAGA  TGAAGATAAA  TTATTAGCAC
                                                  TOX F1
  501   TATATTTCTT  TGCTTCACAA  GAACTTCCAA  TGGAGGCAAA  TCAACAATCA
                                               ←
                                                    TOX R3
  551   AATGCAGCAA  ATTTTTTTAA  AGTAATTGAT  TTTTTACTTA  TCTTATCTGC

  601   TGTAACATCA  CTGGGAAAAA  GGATTTTTTC  AAAAAATTTT  TACAATGGTC

  651   TAGAAACTAA  ATCATTAGAG  AATTATATTG  AGAGAAAAAA  ACTTTCTAAA
                                                TOX F3
  701   CCTTTCTTTC  GACCACCGCA  GAAGTTACCT  GATGGCAGAA  CAGGCTACTT

  751   GGCCGGTCCA  ACAAAAGCGC  CTAAATTGCC  AACAACGTCT  TCTACAGCAA
                                               ←
                                                 TOX R4
  801   CAACGTCTAC  AGCAGCTTCA  TCTAATTGGA  GAGTTAGTTT  GCAAAAACTT

  851   AGAGATAACC  CATCCAGAAA  TACATTTATG  AAAATGGATG  ATGCTGCAAA

  901   ACGAAAATAT  AGTTCATTTA  TAAAAGAGGT  ACAAAAGGGT  AATGATCCAC

  951   GTGCAGCAGC  AGCAAGTATT  GGTACAAAAA  GCGGCAGTAA  CTTCGAAAAA

 1001   CTGCAAGGTA  GAGATTTATA  TAGTATAAGA  CTAAGCCAAG  AACACAGGGT
                                               ←
                                                 A24AC1
 1051   AACATTCTCC  ATAAATAATA  CTGACCAAAT  AATGGAGATC  CAAAGTGTTG

 1101   GAACTCATTA  CCAAAATATA  |TAA|CCTGATT  TATAGTAGTG  ATAAGACGTA

 1151   AGATAAATAT  GGAAGGTTGT  AATTCTATTG  CACTTCCTCA  GAGGTGACCG

 1201   CTCAG
```

FIGURE 1 .

```
  1  MVIKPVTTPS VIQLTPDDRV TPDDKGEYQP VEKQIAGDII RVLEFKQTNE

 51  SHTGLYGIAY RAKKVIIAYA LAVSGIHNVS QLPEDYYKNK DNTGRIYQEY

101  MSNLLSALLG ENGDQISKDM ANDFTQNELE FGGQRLKNTW DIPDLENKLL

151  EDYSDEDKLL ALYFFASQEL PMEANQQSNA ANFFKVIDFL LILSAVTSLG

201  KRIFSKNFYN GLETKSLENY IERKKLSKPF FRPPQKLPDG RTGYLAGPTK

251  APKLPTTSST ATTSTAASSN WRVSLQKLRD NPSRNTFMKM DDAAKRKYSS

301  FIKEVQKGND PRAAAASIGT KSGSNFEKLQ GRDLYSIRLS QEHRVTFSIN

351  NTDQIMEIQS VGTHYQNI
```
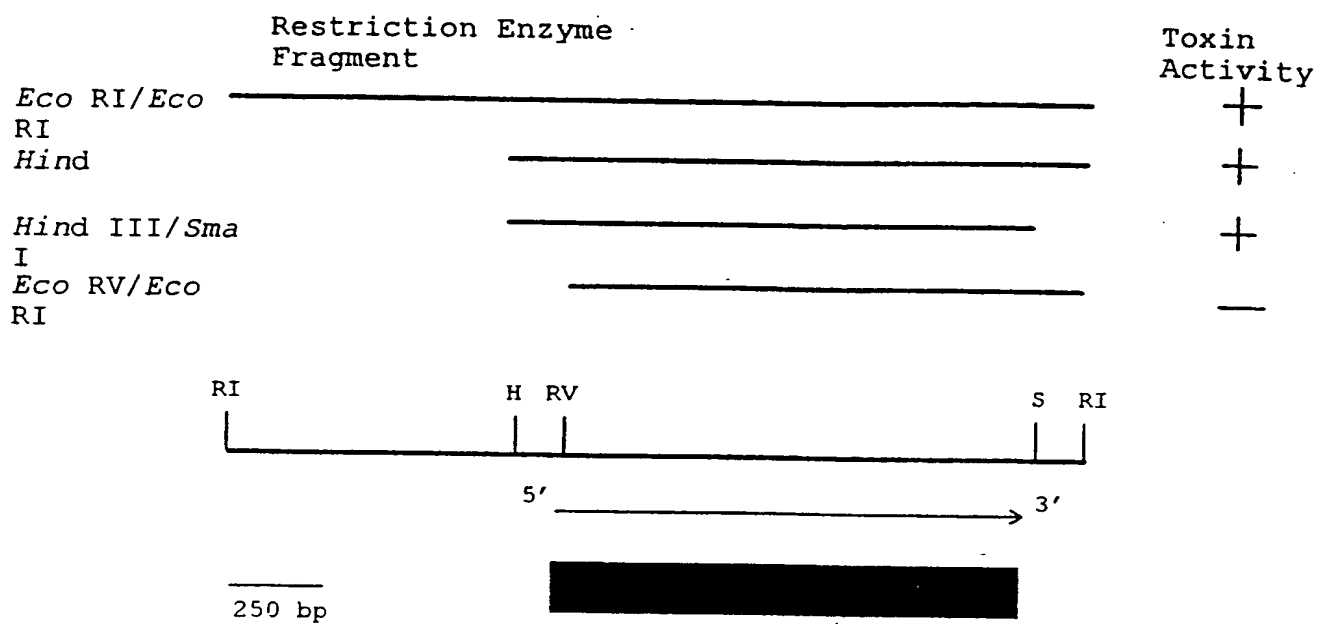
FIGURE 2

Restriction Enzyme
Fragment

Toxin
Activity

Eco RI/Eco
RI ————————————————————————————— +

Hind ——————————————————————— +

Hind III/Sma
I ————————————————————— +

Eco RV/Eco
RI ————————————————— —

RI   H RV   S RI

5' ————————————————————→ 3'

——————
250 bp

FIGURE 3

```
   1  AAGCTTGCTA ATAATTCTTG CGTAAGTTAA TTTTACATTG AAATTAACGC
      Hind III
  51  TTAAAAAGCC AGGGAAAACT CTATATTTAA AGTTGAAATT TATATTAGTA

 101  GCGACAAATT GCGGAGTTTT CTGCCAGAAA TTTCATAGCT CAAATAAACA

 151  TTAACATAAT GGAGAAATAT AATGGTTATA CAATTAACAC CTGATGATAG

 201  AAGTGGATAT CCACCCGTTG AAAAGCAAAT AGCAGGAGAT ATAGTACGTA
          Eco RV
 251  TACTAAACTT TAAGCAAACA GATGAGGGTC ATACAGCATC ATATGGAATT

 301  GAATATCGAG CTAAGAAAAT AATATTAGCT TACGCTTTGG CTGTAAGTGG
                                              AC4R
 351  TATTCATAAT GTATCTAAAC TTCCTGATGA CTATTATAAG AATAAAGAGA

 401  CTGCTGAGAG AATTTATCAA GAATATATGT CTAATCTTTC ATCTGCACTA

 451  TTAGGTGAAA ATGGTGATCA AATTTCTAAA GATATGGCAA ATGGTTTTTA
                       AC2F
 501  TAAGAATGAA CTGGATTTTG AAGGTCAATA TCCTCAAAAC ATTTGGAATG

 551  TTCCTGAGCT TGAAAATAAA CCATTGAGTG CTTATTCAGA TGACGATAAA
                                              AC7R
 601  TTATTAGCAC TATATTTTTT CTCTGTACAG GAAATTCCAC TGGAGGAAAA

 651  TCAACAATCA AATGCCGCAA GATTTTTTAA ATTAATTGAT TTCTTATTTA

 701  CCTTATCTGC TGTAACTTCA CTGGGAAGGA GGATTTTTTC AAAAAACTTT

 751  TACAATGGAT TAGAGGCTAA ATCATTAGAG AATTATATTG AGAGAAAAAA
                            AC6F
 801  ACTTTCTAAA CCTTTCTTTC GACCACCGCA GAGATTACCT GATGGCAGAA

 851  TAGGTTATTT GGCTGGACCA ACAGAAGCGC CTAAATGGAG AGTGAGTTTT
                              AC5R
 901  AAAGAACTTA AAAATAACAA ATCTAGGAAT GGATTTTCTA ATATGGAAGG

 951  GGCTGCAAAA CAAAAGTATA GTTCATTTAT AAAAGAGGTA CAAAAGGGTA

1001  ACGCTCCACA GACAGCAGCG AAAAGTATTG GTACAGCCAG TGGCAGTAAC

1051  CTGGAAAAAT TGCCGAATAA TTTATATAGT GTGAGGCTAA GCCAAAAAGA
                                                      AC3F
1101  CAGGGTAACC TTTACTCAAA ATGATACTGA CAATACAATG ACGGTTCATA
                                                     AC8R
1151  GTGTTGGAAC TCATTATAAA AATATATGAT GAGTAATCTC TGACTTCGAT

1201  TGACAGAGCA TTTTTAAGCT CTCATTTTCT CAACGGGAGT CTCATAAGGC

1251  GTTTTACTTT TCAAGCCACT ATGTGGTCTG TGATAATTGT AAAACGCCTT

1301  CTTTTAGCCA ATACACTTTA CTACCAAGAA AATATATACC CTATGGATTT
                  V16AC1
1351  CAAGATGGAT CGCGGCGGCA AGGGAGCGAA TCCCCGGG
                                        Sma I
```

FIGURE 4

```
  1   MVIQLTPDDR SGYPPVEKQI AGDIVRILNF KQTDEGHTAS YGIEYRAKKI

 51   ILAYALAVSG IHNVSKLPDD YYKNKETAER IYQEYMSNLS SALLGENGDQ

101   ISKDMANGFY KNELDFEGQY PQNIWNVPEL ENKPLSAYSD DDKLLALYFF

151   SVQEIPLEEN QQSNAARFFK LIDFLFTLSA VTSLGRRIFS KNFYNGLEAK

201   SLENYIERKK LSKPFFRPPQ RLPDGRIGYL AGPTEAPKWR VSFKELKNNK

251   SRNGFSNMEG AAKQKYSSFI KEVQKGNAPQ TAAKSIGTAS GSNLEKLPNN

301   LYSVRLSQKD RVTFTQNDTD NTMTVHSVGT HYKNI
```

FIGURE 5

```
 17 ATGGTTATTAAACCCGTAACAACTCCGAGTGTAATACAATTAACGCCTGA  66
                                                      |
172 ..................................................ATGGT 176

 67 TGATAGAGTAACGCCTGATGATAAAGGTGAATATCAACCCGTTGAAAAGC 116
    |      | |||| ||||||||||| | ||| |||||| |||||||||||||||
177 TATACAATTAACACCTGATGATAGAAGTGGATATCCACCCGTTGAAAAGC 226

117 AAATAGCGGGAGATATAATACGTGTACTAGAATTCAAGCAAACAAATGAA 166
    ||||||| |||||||||| ||||| ||||| | || ||||||||| ||||
227 AAATAGCAGGAGATATAGTACGTATACTAAACTTTAAGCAAACAGATGAG 276

167 AGTCATACAGGATTGTATGGAATTGCATATCGAGCTAAGAAAGTAATAAT 216
    ||||||||| || |||||||||| |||||||||||||||||| ||||| |
277 GGTCATACAGCATCATATGGAATTGAATATCGAGCTAAGAAAATAATATT 326

217 AGCATATGCTTTAGCGGTAAGTGGTATTCATAATGTCTCTCAACTTCCAG 266
    ||| || ||||| || ||||||||||||||||||||| ||| |||||| |
327 AGCTTACGCTTTGGCTGTAAGTGGTATTCATAATGTATCTAAACTTCCTG 376

267 AAGACTATTATAAAAATAAGGATAACACAGGTAGAATTTATCAAGAATAC 316
    | |||||||||||| ||||| || |    | |  ||||||||||||||||||
377 ATGACTATTATAAGAATAAAGAGACTGCTGAGAGAATTTATCAAGAATAT 426

317 ATGTCTAATCTTTTATCTGCACTATTGGGTGAGAATGGTGATCAAATTTC 366
    |||||||||||| |||||||||| ||||| ||||||||||||||||||||
427 ATGTCTAATCTTTCATCTGCACTATTAGGTGAAAATGGTGATCAAATTTC 476

367 TAAAGATATGGCAAATGATTTTACCCAGAACGAACTGGAGTTTGGAGGTC 416
    ||||||||||||||||||| ||||    |||| ||||||||| ||| |||||
477 TAAAGATATGGCAAATGGTTTTTATAAGAATGAACTGGATTTTGAAGGTC 526

417 AACGTCTTAAAAATACCTGGGATATTCCTGATCTTGAGAATAAACTATTG 466
    ||  || | ||||  |  ||| ||| || |||||||| |||||| |||| |||
527 AATATCCTCAAAACATTTGGAATGTTCCTGAGCTTGAAAATAAACCATTG 576

467 GAAGATTATTCAGATGAAGATAAATTATTAGCACTATATTTCTTTGCTTC 516
    | |||||||||||| |||||||||||||||||||||||||||| || ||
577 AGTGCTTATTCAGATGACGATAAATTATTAGCACTATATTTTTTCTCTGT 626

517 ACAAGAACTTCCAATGGAGGCAAATCAACAATCAAATGCAGCAAATTTTT 566
    ||| ||| |||||| |||||| ||||||||||||||||||||| |||| ||||
627 ACAGGAAATTCCACTGGAGGAAAATCAACAATCAAATGCCGCAAGATTTT 676

567 TTAAAGTAATTGATTTTTTACTTATCTTATCTGCTGTAACATCACTGGGA 616
    ||||| ||||||||||| ||| ||| |||||||||||||||| |||||||||
677 TTAAATTAATTGATTTCTTATTTACCTTATCTGCTGTAACTTCACTGGGA 726

617 AAAAGGATTTTTTCAAAAAATTTTTACAATGGTCTAGAAACTAAATCATT 666
    |   ||||||||||||||||| ||||||||||||   |||| |||||||||||
727 AGGAGGATTTTTTCAAAAAACTTTTACAATGGATTAGAGGCTAAATCATT 776

667 AGAGAATTATATTGAGAGAAAAAAACTTTCTAAACCTTTCTTTCGACCAC 716
    ||||||||||||||||||||||||||||||||||||||||||||||||||
777 AGAGAATTATATTGAGAGAAAAAAACTTTCTAAACCTTTCTTTCGACCAC 826

717 CGCAGAAGTTACCTGATGGCAGAACAGGCTACTTGGCCGGTCCAACAAAA 766
    ||||||| |||||||||||||||||| ||| || ||||| || |||||| ||
827 CGCAGAGATTACCTGATGGCAGAATAGGTTATTTGGCTGGACCAACAGAA 876
```

Figure 6

```
 767 GCGCCTAAATTGCCAACAACGTCTTCTACAGCAACAACGTCTACAGCAGC 816
     ||||||||||
 877 GCGCCTAAA......................................... 885

 817 TTCATCTAATTGGAGAGTTAGTTTGCAAAAACTTAGAGATAACCCATCCA 866
                |||||||| ||||| || |||||| | ||||| ||| |
 886 ..........TGGAGAGTGAGTTTTAAAGAACTTAAAAATAACAAATCTA 925

 867 GAAATACATTTATGAAAATGGATGATGCTGCAAAACGAAAATATAGTTCA 916
     | ||| |||| || |||||| | |||||||||| ||| ||||||||||
 926 GGAATGGATTTTCTAATATGGAAGGGGCTGCAAAACAAAAGTATAGTTCA 975

 917 TTTATAAAAGAGGTACAAAAGGGTAATGATCCACGTGCAGCAGCAGCAAG 966
     |||||||||||||||||||||||||||||| | ||||| ||||||| |||
 976 TTTATAAAAGAGGTACAAAAGGGTAACGCTCCACAGACAGCAGCGAAAAG 1025

 967 TATTGGTACAAAAAGCGGCAGTAACTTCGAAAAACTGCAAGGTAGAGATT 1016
     |||||||||| || |||||||||| | |||||| ||| || |||
1026 TATTGGTACAGCCAGTGGCAGTAACCTGGAAAAATTGCCGAATA...ATT 1072

1017 TATATAGTATAAGACTAAGCCAAGAACACAGGGTAACATTCTCCATAAAT 1066
     |||||||| | || |||||||||| || |||||||||| || | ||||
1073 TATATAGTGTGAGGCTAAGCCAAAAAGACAGGGTAACCTTTACTCAAAAT 1122

1067 AATACTGACCAAATAATGGAGATCCAAAGTGTTGGAACTCATTACCAAAA 1116
     |||||||| | | |||| | | || |||||||||||||||||| ||||
1123 GATACTGACAATACAATGACGGTTCATAGTGTTGGAACTCATTATAAAAA 1172

1117 TATA... 1120
     ||||
1173 TATATGA 1179
```

Figure 6 continued

```
  1 MVIKPVTTPSVIQLTPDDRVTPDDKGEYQPVEKQIAGDIIRVLEFKQTNE 50
             ||||||||||         ::|.||||||||||||:|:|:||||:|
  1 .........MVIQLTPDDR......SGYPPVEKQIAGDIVRILNFKQTDE 35

 51 SHTGLYGIAYRAKKVIIAYALAVSGIHNVSQLPEDYYKNKDNTGRIYQEY 100
    :||:  |||.|||||:|:||||||||||||||.||:||||||:..:|||||
 36 GHTASYGIEYRAKKIILAYALAVSGIHNVSKLPDDYYKNKETAERIYQEY 85

101 MSNLLSALLGENGDQISKDMANDFTQNELEFGGQRLKNTWDIPDLENKLL 150
    ||||  ||||||||||||||||||||||:|  .|||:|:||   .|.|::|:||||  |
 86 MSNLSSALLGENGDQISKDMANGFYKNELDFEGQYPQNIWNVPELENKPL 135

151 EDYSDEDKLLALYFFASQELPMEANQQSNAANFFKVIDFLLILSAVTSLG 200
    ..|||:|||||||||.  ||:|:|.||||||||.|||:||||:.|||||||
136 SAYSDDDKLLALYFFSVQEIPLEENQQSNAARFFKLIDFLFTLSAVTSLG 185

201 KRIFSKNFYNGLETKSLENYIERKKLSKPFFRPPQKLPDGRTGYLAGPTK 250
    :|||||||||||||.|||||||||||||||||||||:|||||.|||||||.
186 RRIFSKNFYNGLEAKSLENYIERKKLSKPFFRPPQRLPDGRIGYLAGPTE 235

251 APKLPTTSSTATTSTAASSNWRVSLQKLRDNPSRNTFMKMDDAAKRKYSS 300
    |||                  ||||:..|::|.|||.|  .|::|||.||||
236 APK................WRVSFKELKNNKSRNGFSNMEGAAKQKYSS 268

301 FIKEVQKGNDPRAAAASIGTKSGSNFEKLQGRDLYSIRLSQEHRVTFSIN 350
    |||||||||.|..||  ||||  ||||:|||..  :|||:||||..||||. |
269 FIKEVQKGNAPQTAAKSIGTASGSNLEKLPN.NLYSVRLSQKDRVTFTQN 317

351 NTDQIMEIQSVGTHYQNI 368
    :||..|.::||||||.||
318 DTDNTMTVHSVGTHYKNI 335
```
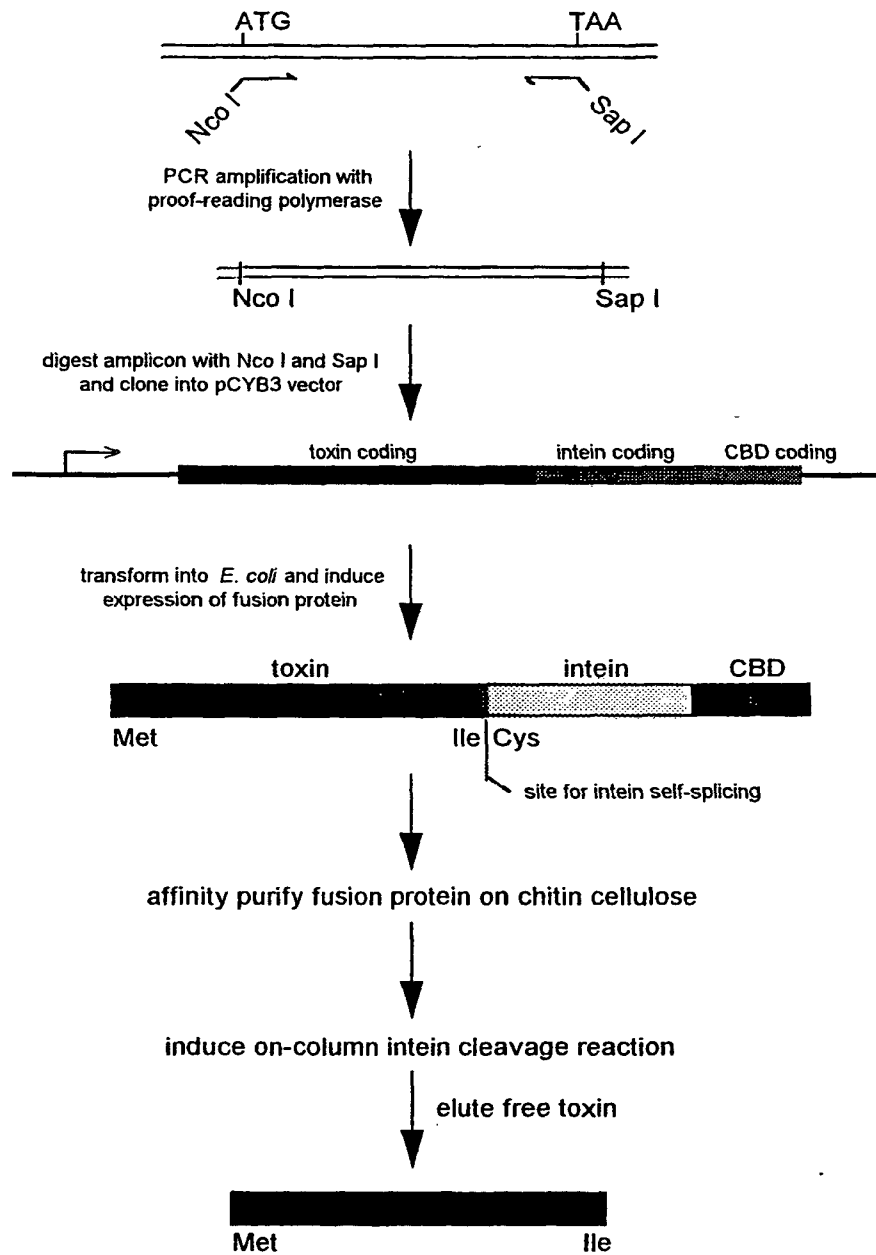
FIGURE 7

ATG   TAA

PCR amplification with
proof-reading polymerase

Nco I   Sap I

digest amplicon with Nco I and Sap I
and clone into pCYB3 vector

toxin coding   intein coding   CBD coding

transform into *E. coli* and induce
expression of fusion protein

toxin   intein   CBD
Met   Ile Cys
                      site for intein self-splicing

affinity purify fusion protein on chitin cellulose

induce on-column intein cleavage reaction

elute free toxin

Met   Ile

Figure 8.